Advancements in the DOE Joint Genome Institute's High Throughput Production Sequencing Program

Susan M. Lucas, John C. Detter , Tijana Glavina, Nancy Hammon, Sanjay Israni, Martin Pollard, Alex Copeland, Kerrie Barry, Simon Roberts, Feng Chen, Nathaniel Slater, Samuel Pitluck, Christopher Daum, Paul Richardson, Eddy Rubin, and the JGI Sequencing Team

U.S. DOE Joint Genome Institute, Walnut Creek, CA 94598

The Department of Energy's (DOE) Joint Genome Institute (JGI) Production Genomics Facility (PGF) is responsible for high throughput sequencing. The sequencing process is divided into three subgroups; Library Support, Sequencing Prep and Capillary Electrophoresis, which collectively transform a variety of input DNAs into high quality shotgun sequence. Transformation stocks from whole genome shotgun libraries enter the process at the Library Support step, where they are plated, picked and the 3kb and 8kb libraries are sent on for template preparation using Templiphi. Fosmid DNA is prepared in parallel using the SPRInt protocol from Agencourt. Clones are then end sequenced using Big Dye Terminator or Dyenamic ET chemistry kits and run on their respective platforms – ABI 3730xl or MegaBACE 4000. A series of automated post-sequencing data processing steps then convert the raw shotgun sequence into assembled contigs, where the data is QC'd and prepared for release. In our efforts to scale production over the last year, the process has undergone dynamic changes to increase throughput and efficiency. Changes have occurred in sequencing chemistry to both reduce cost as well as generate high quality sequence for GC-rich templates. Tracking efficiency has increased due to the release of a new LIMS. The implementation of a training program and preventative maintenance program has allowed for stability of both instruments and staff. The conversion of the MegaBACE 4000 to the MegaBACE 4500 has had a direct effect on increased readlenths and pass rates. In combination, all of these changes have resulted in significant improvements in pass rates, readlengths, stability and cost savings, enabling the PGF to put through several large sequencing projects through including Xenopus tropicalis, Nematostella vectensis, Emiliania huxyeli, and over fifty microbes, maintaining a monthly throughput of 3.9 million lanes resulting in ~2.4 billion Q20 base pairs.

This work was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, the Lawrence Berkeley National Laboratory under contract No. DE-AC03-76SF00098, and the Los Alamos National Laboratory under contract No. W-7405-ENG-36.